Course OT - Exo #2

Exo 1: Define for $a \in \mathbb{R}_+^n$, $b \in \mathbb{R}_+^m$, $\lambda \in [0, \min(\langle a, 1\rangle, \langle b, 1\rangle)]$ the partial OT pb:

$$\min_{P \in \mathbb{R}_+^{n\times m}} \{ \langle C, P\rangle : P\mathbb{1} \leq a,\ P^T\mathbb{1} \leq b,\ \langle P, \mathbb{1}\mathbb{1}^T\rangle = \lambda \}$$

Compute the dual problem

Exo 2: We consider the following unbalanced OT pbm:

$$(\mathrm{UOT}_\tau) \quad \min_{P \in \mathbb{R}_+^{n\times m}} \langle C, P\rangle + \tau\, \mathrm{KL}(P\mathbb{1}|a) + \tau\, \mathrm{KL}(P^T\mathbb{1}|b) \triangleq f_\tau(P)$$

Compute the dual to this pb.

Exo 3: We consider the entropic regul$^\circ$:

$$\min_{P \in \mathbb{R}_+^{n\times m}} f_\tau(P) + \varepsilon H(P) \triangleq f_\tau^\varepsilon(P)$$

Compute the dual problem. Deduce "generalized" Sinkhorn iteration to minimize f_τ^ε

Exo 4: One can show that for $\tau \to 0$,

$$\min_P \frac{1}{\tau} f_\tau(P) \longrightarrow \min_{P \in Z} \mathrm{KL}(P\mathbb{1}|a) + \mathrm{KL}(P^T\mathbb{1}|b)$$

$$Z \triangleq \operatorname{Argmin}_{P \geq 0} \langle P, C\rangle$$

Compute this value when $C = d^p$ with d a distance

$n = m$

Exo 5: Consider Schrodinger problem:

$$\min_{\pi} \left\{ \int c(x,y)\, d\pi(x,y) + \varepsilon KL(\pi | \alpha\otimes\beta) : \begin{array}{l} \pi_1 = \alpha \\ \pi_2 = \beta \end{array} \right\}.$$

Show that its solution is of the form: $\frac{d\pi}{d\alpha\otimes\beta}(x,y) = \underbrace{e^{-\frac{c(x,y)}{\varepsilon}}}_{\triangleq K(x,y)} \cdot u(x)\cdot v(y)$

and that:

$$v(y) = \underbrace{\frac{1}{\int K(x,y)\, u(x)\, d\alpha(x)}}_{\triangleq S_\alpha(u)(y)} \quad , \quad u(y) = \underbrace{\frac{1}{\int K(x,y)\, v(y)\, d\beta(y)}}_{\triangleq S_\beta(v)(x)}$$

Exo 6: Assume $c(x,y) = |x-y|^2$ and that (α, β) are gaussians. Show that if $u(x) = \exp(A|x-m|^2)$ for some A and m, then $S_\beta(u)$ is of the same form.

Hint: use that convol° of gaussians is gaussians.

Deduce that the sol° of Schrodinger problem π is a Gaussian.

Exo 7: If $c(x,y) = -\langle x,y\rangle$ show that sinkhorn iterations in log domain:

$$\begin{cases} f(x) \leftarrow -\varepsilon \log \int \exp\left(\frac{-c(x,y) + g(y)}{\varepsilon}\right) d\beta(y) \\ g(y) \leftarrow -\varepsilon \log \int \exp\left(\frac{-c(x,y) + f(x)}{\varepsilon}\right) d\alpha(x). \end{cases}$$

Defines concave functions.

Exo 8: ~~Show that during those iteration on always have~~

$$\cancel{f(x) + g(y) \leqslant c(x,y)}$$

- ~~Deduce a lower bound on the OT cost during the iterations.~~
- ~~Explain why the Sinkhorn iteration can be stabilized as:~~

$$\cancel{f(x) \leftarrow -\varepsilon \log \int \exp\left(\frac{-c(x,y) + f(x) + g(y)}{\varepsilon}\right) d\beta(y) + f(x)}$$

Correct^{c} Exo 1 : this is a linear prog.

The coupling $P \triangleq \frac{a \otimes b}{|a|_1 \cdot |b|_1} \cdot \lambda$ satisfy the constraints

$\rightsquigarrow$ One can exchange inf/sup = sup/inf

$$\inf_{P \geq 0} \{ \langle C, P \rangle : \underset{\substack{\uparrow \\ f \in \mathbb{R}^n_-}}{a - P\mathbb{1} \geq 0}, \underset{\substack{\uparrow \\ g \in \mathbb{R}^m_-}}{b - P^T\mathbb{1} \geq 0}, \underset{\substack{\uparrow \\ s \in \mathbb{R}}}{\langle P, \mathbb{1}\mathbb{1}^T \rangle = \lambda} \}.$$

$$\inf_{P \geq 0} \sup_{f \leq 0, g \leq 0} \langle C, P \rangle + \langle a - P\mathbb{1}, f \rangle + \langle b - P^T\mathbb{1}, g \rangle + (-\langle P, \mathbb{1} \rangle + \lambda) s$$

$$\sup_{f \leq 0, g \leq 0, s} (\langle a, f \rangle + \langle b, g \rangle + \lambda s) + \inf_{P \geq 0} \langle P, C - f \oplus g - s \rangle$$

$$\sup_{f g s} \{ \langle a, f \rangle + \langle b, g \rangle + \lambda s : \begin{matrix} f \leq 0 \\ g \leq 0 \end{matrix}, f \oplus g + s \leq C \}.$$

Cor Exo 2 : Write $KL(\mu | a) = \sup_f \langle f, \mu \rangle - KL^*(f | a)$

where $KL^*(f | a) \underset{\circledast}{=} \sup_\mu \langle f, \mu \rangle - KL(\mu | a)$

optimality in $\circledast$: $f - \log(\mu / a) = 0 \Rightarrow \mu = a \odot e^f$

$$\rightsquigarrow KL^*(f | a) = \langle f, \mu \rangle - \langle \mu, \log\left(\frac{\mu}{a}\right) - 1 \rangle - \langle a, 1 \rangle = \langle f, \mu \rangle - \langle \mu, f - 1 \rangle - \langle a, 1 \rangle$$

$$= \langle \mu, 1 \rangle - \langle a, 1 \rangle = \langle a, e^f - 1 \rangle = \sum_i a_i (e^{f_i} - 1)$$

$$\inf_{P\geq 0} \langle C,P\rangle + \tau KL(P1|a) + \tau KL(P^T1|b)$$

$$= \inf_{P\geq 0} \sup_{f,g} \langle C,P\rangle + \tau\left(\langle P1, f\rangle + \langle P^T1, g\rangle - KL^*(f|a) - KL^*(g|b)\right)$$

$$= \sup_{f,g} -\tau KL^*(f|a) - \tau KL^*(g|b) + \inf_{P\geq 0} \langle C + \tau f \oplus g, P\rangle$$

$$= \sup \left\{ -\tau\left(\sum_i a_i(e^{f_i}-1) + \sum_j b_j(e^{g_j}-1)\right) : C + \tau f\oplus g \geq 0\right\}$$

$$= \sup_{f,g} \left\{ \sum_i a_i \underbrace{\tau(1-e^{-f_i/\tau})}_{\substack{\tau\to+\infty\\ \sim f_i}} + \sum_j b_j \underbrace{\tau(1-e^{-g_j/\tau})}_{\sim g_j} : f\oplus g \leq C\right\} \quad \tau f \to -f$$

Cor
Exo 3 : Same comput$^{\underline{c}}$ ...

$$\sup_{f,g} \tau\langle a, 1-e^{-f/\tau}\rangle + \tau\langle b, e^{-g/\tau}\rangle + \varepsilon\sum_{ij}\exp\left(\frac{C_{ij}+f_i+g_j}{\varepsilon}\right)$$

Cor
Exo 4 :

$$\min_{P\in Z} KL(P1|a) + KL(P^T1|b) \quad (*)$$

$$Z \triangleq \operatorname{Argmin}_{P\geq 0} \langle P, d^P\rangle = \{P = \operatorname{diag}(q) : q\in\mathbb{R}^n_+\}$$

$$(*) \quad \min_{q\in\mathbb{R}^n_+} KL(q|a) + KL(q|b) = E(q)$$

At optimality: $\log(q/a) + \log(q/b) = 0 \iff q = \sqrt{a \odot b}$

$$E(q) = \sum_i \underbrace{q_i \log(q_i/a_i) + q_i \log(q_i/b_i)}_{=0} + a_i - q_i + b_i - q_i$$

$$= \sum_i a_i + b_i - 2\sqrt{a_i b_i} = \sum_i (\sqrt{a_i} - \sqrt{b_i})^2 = \|\sqrt{a} - \sqrt{b}\|_2^2$$

(squared Hellinger distance)

Cor

<u>Exo 5</u>: $\inf_{\pi \geq 0} \left\{ \int c \, d\pi + \varepsilon KL(\pi | \alpha \otimes \beta) \;\; \begin{matrix} \pi_1 = \alpha \\ \pi_2 = \beta \end{matrix} \right\}$. Let's compute the dual:

$$= \inf_{\pi \geq 0} \sup_{f,g \in C} \left\{ \int c \, d\pi + \varepsilon KL(\pi | \alpha \otimes \beta) + \langle \alpha - \pi_1, f \rangle + \langle \beta - \pi_2, g \rangle \right\}$$

$$= \sup_{f,g} \int f \, d\alpha + \int g \, d\beta + \underbrace{\inf_{\pi} \langle \pi, C - f \oplus g \rangle + \varepsilon KL(\pi | \alpha \otimes \beta) + \iota_{\geq 0}(\pi)}$$

$$A = -\varepsilon KL^*\left(\frac{-C + f \oplus g}{\varepsilon} \Big| \alpha \otimes \beta\right)$$

<u>Lem</u>: $KL^*(H | \alpha \otimes \beta) = \int \left(e^{H(x,y)} - 1\right) d\alpha \otimes \beta(x,y)$

(see exo 2 in discrete)

$$= \sup_{f,g} \int f \, d\alpha + \int g \, d\beta - \varepsilon \iint \exp\left(\frac{-c(x,y) + f(x) + g(y)}{\varepsilon}\right) d\alpha(x) \, d\beta(y) + \varepsilon$$

Primal/Dual Rel$^{\circ}$: $\partial / \pi = 0$ in (A)

$$C - f \oplus g + \varepsilon \log\left(\frac{d\pi}{d\alpha \otimes \beta}\right) = 0 \Rightarrow \frac{d\pi}{d\alpha} = \exp\left(\frac{-C + f \oplus g}{\varepsilon}\right)$$

ie. $\frac{d\pi}{d\alpha \otimes \beta}(x,y) = u(x)\, v(y)\, K(x,y)$ with $\begin{cases} u(x) = \exp(f(x)/\varepsilon) \\ v(y) = \exp(g(y)/\varepsilon) \\ K = \exp(-C/\varepsilon) \end{cases}$

Sinkhorn Algo: alternate optim$^{\circ}$ // f then g.

$\partial / f = 0$ imply: "$\alpha(x) - \int \exp\left(\frac{-c(x,y) + f(x) + g(y)}{\varepsilon}\right) d\beta(y)\, \alpha(x) = 0$"

ie $\quad \underbrace{\exp\left(\frac{f(x)}{\varepsilon}\right)}_{u(x)} \cdot \int \exp\left(\frac{g(y) - c(x,y)}{\varepsilon}\right) d\beta(y) = 1$ "

$$u = \frac{1}{K_\beta v} \qquad K_\beta(v) \stackrel{\Delta}{=} \int K(x,y)\, v(y)\, d\beta(y)$$

log-sum-exp format :

$$f(x) = -\varepsilon \log \int \exp\left(\frac{-C(x,y) + g(y)}{\varepsilon}\right) d\beta(y)$$

$$= \text{"} \min{}^{\beta}_{\varepsilon}\left(C(x,y) - g(y)\right) \text{"} \quad \text{soft min}$$

$$\min{}^{\beta}_{\varepsilon}(h) = -\varepsilon \log \int \exp\left(-h(y)/\varepsilon\right) d\beta(y).$$

Similarly $\quad g(y) = \min{}^{\alpha}_{\varepsilon}\left(C(x,y) - f(x)\right) . \qquad \min{}^{\alpha}_{\varepsilon} \xrightarrow{\varepsilon \to 0} \min$

Stabilisat$^{\circ}$: $\quad \min{}^{\alpha}_{\varepsilon}(h) = \min{}_{\varepsilon}\left(h - \min(h)\right) + \min(h)$.

Cor
Exo 6 : wlog one can assume means are 0

$$G_\sigma(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{|x|^2}{2\sigma^2}\right) \qquad \frac{d\alpha}{dx} = G_{\sigma_\alpha} \qquad \frac{d\beta}{dx} = G_{\sigma_\beta} \qquad k(x,y) \propto G_\varepsilon(x-y)$$

$\varepsilon \to 2\varepsilon^2$

$$S_\alpha(\mu) = \left[\int k(x,y)\underbrace{\mu}_{=G_{\sigma_\mu}}(x)\,d\alpha(x)\right]^{-1} \propto \left[G_\varepsilon * \left(G_{\sigma_\mu} \odot G_{\sigma_\alpha}\right)\right]^{-1}$$

Lem : $G_s \odot G_t \propto G_r \qquad \frac{1}{s} + \frac{1}{t} = \frac{1}{r}$

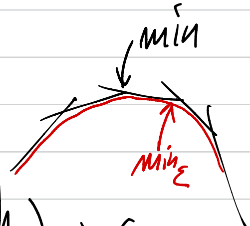
$G_s * G_t = G_r \qquad r = \sqrt{s^2 + t^2}$ ...

$1/G_s \propto G_{-s}$

$\Rightarrow S_\alpha(G_s) = G_t$ with $t = -\left[\sqrt{\varepsilon^2 + \left(\frac{1}{1/s + 1/\sigma_\alpha}\right)^2}\right]$

$\Rightarrow$ solve for $s = t$ ...

Cor
Exo 7 : $f = \min^\beta_\varepsilon\left(-\langle y, \cdot\rangle - g(y)\right)$



Lemma : if h_y are concave, $\min^\beta_\varepsilon(h_y)$ is concave

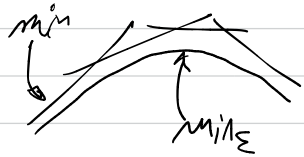
Proof : see example 3.14 in Boyd's Book

Cor

Exo 8 : one has

This is not correct

$$f(x) = \min^{\beta}_{\varepsilon}\big(c(x,\cdot) - g\big)$$



Lem : $\min^{\beta}_{\varepsilon}(h) \leq \min(h)$

Proof : $\min_{\varepsilon}(h) - \min(h) = \underbrace{-\varepsilon \log \underbrace{\int \underbrace{\exp\Big(\underbrace{\underbrace{\tfrac{-h(x)+\min(h)}{\varepsilon}}}_{\geq 0}\Big)}_{\geq 1} d\alpha(x)}_{\geq 1}}_{\leq 0}$

$$\Rightarrow f(x) \leq c(x,y) - g(y) \quad \forall y$$

- this imply that (f,g) is dual feasible, hence

$$OT(\alpha,\beta) \geq \int f\, d\alpha + \int g\, d\beta$$

- If f is an iterate of Sinkhorn, $f \oplus g \leq c$ and thus

$$g(y) \leftarrow \min_{\varepsilon}\big(c(\cdot,y) - f\big) \quad \leftarrow \text{unstable}$$

$$= \min_{\varepsilon}\big(\underbrace{c(\cdot,y) - f - g(y)}_{\geq 0}\big) + g(y) \quad \text{stable}$$